Cut-off and Escape Behaviors
for Birth and Death Chains on Trees

O. Bertoncini

February 11, 2011

Abstract

We consider families of discrete time birth and death chains on trees, and show that in presence of a drift toward the root of the tree, the chains exhibit cut-off behavior along the drift and escape behavior in the opposite direction.

1   Introduction

Although they originally come from different research fields and seem apparently to be very different phenomena, cut-off and escape behaviors have been related at the level of hitting times for birth and death chains on the line (see [5, 6, 7]). Cut-off behavior refers to the famous cut-off phenomenon first discovered and studied by Aldous and Diaconis in the 80's (see [1, 2, 3]), which is characterized by an almost deterministic asymptotical abrupt convergence to equilibrium for families of Markov processes. In opposition, escape behavior is usually associated to the exit from metastability for a system trapped in local minimum of the energy profile. In that case, the transition to equilibrium occurs at "unpredictable" exponential times (see [8]).

We refer to the introduction in [5] for a comparative historical discussion of the phenomena. In this reference —dealing with birth-and-death chains on the line with drift toward the origin— cut-off and escape phenomena are characterized by their distinct hitting-time behavior. It is shown that, under suitable drift hypotheses, the chains exhibit both cut-off behavior toward zero and escape behavior for excursions in the opposite direction. Furthermore, as the evolutions are reversible, the law of the final escape trajectory coincides with the time reverse of the law of cut-off paths.

In the present work, we extend this study to birth-and-death chains on trees with drift toward the root of the tree. Our chains are the discrete time counterpart of the birth-and-death processes studied by Martínez and Ycart in [13]. Under a "uniform" drift condition, they prove that hitting times of zero starting from any state a (denoted by $T_{a \to 0}$) exhibit cut-off at mean times when a goes to infinity (Theorem 5.1 in [13]). Their drift condition is expressed as an exponential decaying tail of the invariant probability measure π, in the sense that, for all a, the quantity $\pi(B_a)/\pi(a)$ is uniformly bounded by some K. In comparison, our drift condition concerns only the branch which contains the state a (see formulas (15) and (17)), and we allow the upper bound K_a to go to infinity, as long as (16) remains valid (see Section 2.4). With this assumption, we can prove the cut-off behavior of $T_{a \to 0}$ (Proposition 2.2) and that the typical time scale $\mathbb{E}[T_{a \to 0}]$ of this convergence is negligible compare to the mean escape time from zero to a (Proposition 2.3). If in addition we have a control on trajectories outside the branch which contains a (Conditions (19)), we get the escape behavior of $T_{0 \to a}$ (Theorem 2.4).

Note that until the drift condition is satisfied, our results apply to any tree. No particular assumptions are made on the degrees which can be non-homogeneous and non-bounded.

1

The paper is organized as follows. We first recall the general definition of the two types of behaviors in Section 1.1. Then, we define our birth-and-death chains model on trees (Sections 2.1, 2.2 and 2.3), and the strong drift toward the root in Section 2.4. The main results are given in Section 2.5 and we discuss the basic example of regular trees in Section 2.6. Section 3 contains exact expressions for the mean hitting times and their second moments. Finally, the main results are proven in Section 4.

1.1 Cut-off and escape behaviors

Cut-off and escape behaviors, will be studied at the level of hitting times. Both types of behavior are asymptotic, in the sense that they are characterized by what happens when a certain parameter a diverges. Let us recall here the relevant definitions.

Definition 1.1. (i) A family of random variables $U^{(a)}$ *exhibits cut-off behavior at mean times* if

$$\frac{U^{(a)}}{\mathbb{E}\left[U^{(a)}\right]} \xrightarrow[a\to\infty]{\text{Proba}} 1 \, . \tag{1}$$

[equivalently, $\lim_{a\to\infty} \mathbb{P}\big(U^{(a)} > c\,\mathbb{E}[U^{(a)}]\big) = 1$ for $c < 1$ and 0 for $c > 1$].

(ii) A family of random variables $V^{(a)}$ *exhibits escape-time behavior at mean times* if

$$\frac{V^{(a)}}{\mathbb{E}\left[V^{(a)}\right]} \xrightarrow[a\to\infty]{\mathcal{L}} \exp(1) \, . \tag{2}$$

We refer to [5] for a discussion about the motivations of these definitions, as well as general sufficient conditions for them to occur.

2 Model and results

In order to study cut-off and escape behaviors, we will consider families of random walks $X^{(a)}(t)$ defined on a tree I_a. Each of the chains we are defining below is the discrete time counterpart of the birth-and-death processes on trees studied in [13].

2.1 The trees

A tree is an undirected connected graph $G = (I, E)$ with no "nontrivial closed loop", where I is the set of vertices (or nodes) and E the set of edges (or links).
We define a partial order $\preceq$ on I by choosing a node (denoted by 0) to be the root of the tree: for $x, y \in I$, we say that x is before y ($x \preceq y$) if $x = y$ or if it exists a path from y to zero containing x. Thus each element x of the tree (except the root) has a unique parent denoted by $p(x)$:

$$\forall x \in J \, , \ \exists! \, p(x) \ s.t. \quad p(x) \preceq x \quad and \quad (p(x), x) \in E \, , \tag{3}$$

where $J := I \setminus \{0\}$.
There also exists a unique path from x to zero, and we denote by $d(x)$ its length (depth) and by $\ell(x)$ the set of vertices on that path (except 0):

$$\forall x \in J \, , \ \ell(x) := \{x_0, x_1, \ldots, x_{d(x)-1}\} \, , \tag{4}$$

where $x_i := p^i(x)$ for $i = 0, \ldots, d(x)-1$. Note that according to this notation $x_0 = x$ and $x_{d(x)} = 0$. For the sake of notations, we denote by $\alpha(x)$ the last state before zero on that path: $\alpha(x) = x_{d(x)-1}$. Let us also define $\sigma(x)$ the set of the children of x, and B_x the branch stemming from x:

$$\sigma(x) := \{y : x = p(y)\} \, , \ \text{ and } \ B_x := \{y : x \preceq y\} \, . \tag{5}$$

In Section 3, we will make use of the following decompositions, for any $k = 0, \ldots, d(a)$, of the branch B_{a_k} stemming from the k-th parent $a_k = p^k(a)$ of some $a \in J$, and its complementary $\overline{B}_{a_k}$ (see Figure 1):

$$B_{a_k} = B_a \cup \left(\bigcup_{l=1}^{k} \left(\{a_l\} \bigcup_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} B_c \right) \right), \tag{6}$$

$$\overline{B}_{a_k} = \bigcup_{l=k+1}^{d(a)} \left(\{a_l\} \bigcup_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} B_c \right). \tag{7}$$

We also need to define the set C_{a_k}

$$C_{a_k} := \bigcup_{l=k+1}^{d(a)} \left(\bigcup_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} B_c \right). \tag{8}$$

Note that C_{a_k} is the complementary of the path from a_k to zero in $\overline{B}_{a_k}$: $C_{a_k} = \overline{B}_{a_k} \setminus \bigcup_{l=k+1}^{d(a)} \{a_l\}$.
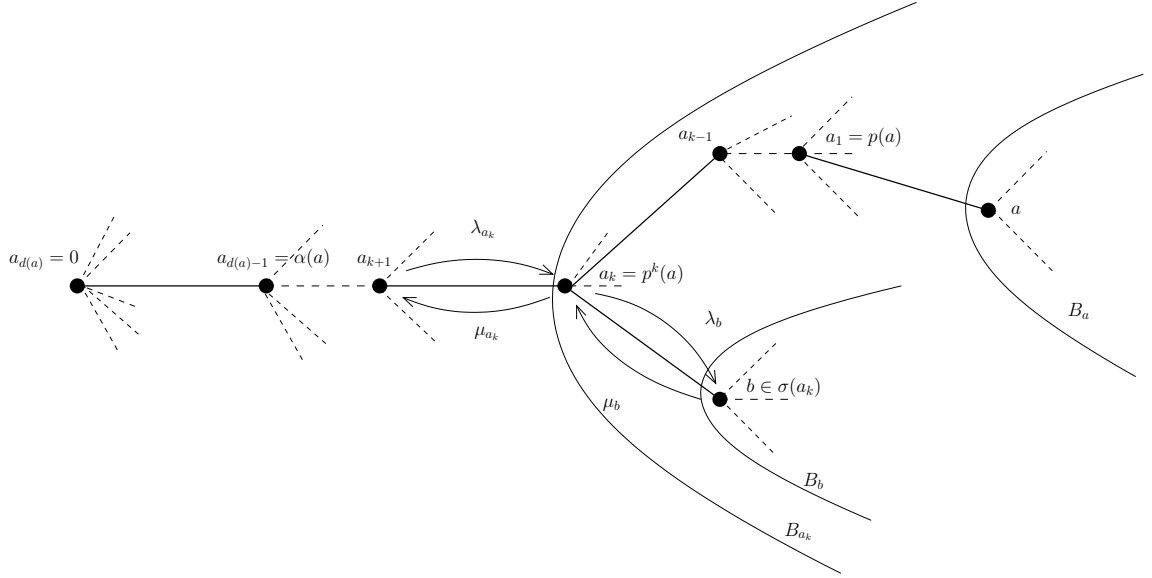

Figure 1: The path from a to zero: sites, transitions and branches

2.2 The chains

We consider irreducible discrete time birth and death chains $X(t)$ on I defined by transition probabilities lx "from parent to children" and μ_x "from children to parent" for each $x \in J$. That is $lx := P(p(x), x)$ and $\mu_x := P(x, p(x))$, where $P(x, y) := \mathbb{P}\big(X(t+1) = y \mid X(t) = x\big)$. The li's and μ_i's sum to 1 in each site:

$$\mu_x + \sum_{y \in \sigma(x)} ly = 1, \quad \forall x \in J, \quad \text{and} \quad \sum_{x \in \sigma(0)} lx = 1, \tag{9}$$

and since the chain is irreducible they are supposed to be non-null.
Note that, as we will consider families of chains $X^{(a)}(t)$, the λ's and μ's may vary with a.

3

As it doesn't change our results and computations, we can also add (at least one) non-null "waiting probabilities" $\kappa_x := P(x,x)$ in order to make the chain aperiodic. In that case, equations (9) become $\mu_x + \kappa_x + \sum_{y \in \sigma(x)} ly = 1$, for all $x \in J$, and $\kappa_0 + \sum_{x \in \sigma(0)} lx = 1$.

If the tree is finite, then the chain is positive recurrent and admits a unique reversible invariant probability measure (i.p.m.) π, which can be expressed for each $x \in J$ in terms of the transition probabilities

$$\pi(x) = \pi(0) \prod_{y \in \ell(x)} \frac{ly}{\mu_y} . \tag{10}$$

This expression is a consequence of the reversibility of π, which ensures that for each $x \in J$ we have $\pi(p(x)) lx = \pi(x) \mu_x$. Then, iteration along the path $\ell(x)$ from x to zero gives (10), where the normalization constant $\pi(0)$ is chosen such that $\sum_{x \in I} \pi(x) = 1$, that is:

$$\pi(0) = \frac{1}{\sum_{x \in I} \prod_{y \in \ell(x)} \frac{ly}{\mu_y}} . \tag{11}$$

In the case where the tree is not finite, we have to suppose that the quantity $\sum_{x \in I} \prod_{y \in \ell(x)} \frac{ly}{\mu_y}$ is finite, and then the chain is positive recurrent and (10) gives the unique i.p.m. π.

2.3 Hitting times

As mentioned above cut-off and escape behavior are studied at the level of hitting times.

More precisely, the families of random variables under consideration are the hitting times (also called first passage times in the literature) of zero starting from some $a \in J$ (resp. times from zero to a), denoted by $T_{a \to 0}$ (resp. $T_{0 \to a}$), with

$$T_{x \to y} = \min\{t \geq 0 : X^x(t) = y\} , \tag{12}$$

where the upper index x means that this is the initial state of the chain ($X^x(0) = x$).

Our results rely on exact expressions for the first and second moments of $T_{p^j(a) \to p^n(a)}$, for any $0 \leq j, n \leq d(a)$, which will be given in Section 3.

Let us just mention now, for all $x \in J$, the expression in terms of the invariant measure for the mean time $\mathbb{E}[T_{x \to p(x)}]$ from child to parent (see Lemma 3.1):

$$\mathbb{E}[T_{x \to p(x)}] = \frac{\pi(B_x)}{\mu_x \pi(x)} . \tag{13}$$

This expression is a key ingredient in our computations and allows also to interpret our drift condition as an exponential decay of π (see next section).

2.4 The drift condition

As seen in Section 1.1, cut-off and escape behaviors are defined for families of variables indexed by some diverging parameter. In the present work, we take a state $a \in J$ such that the length $d(a)$ of the path $\ell(a)$ from a to zero tends to infinity to be our asymptotical parameter.

We denote by $a \to \infty$ the limit when $d(a) \to \infty$, and we define by ℓ the path "from zero to infinity" followed by a when $d(a) \to \infty$: $\ell = \lim_{a \to \infty} \ell(a)$.

Let us give now the definition of strong drift which applies to families of previous defined chains $X^{(a)}(t)$ on trees I_a. When it is necessary, we mark with an upper index (a) the quantities corresponding to $X^{(a)}(t)$.

Definition 2.1. The family of birth-and-death chains $X^{(a)}(t)$ has a *strong drift toward 0 along the path ℓ* (we note 0^ℓ-SD) if

(i) the transition probabilities $\mu_x^{(a)}$ satisfy

$$\inf_{b \in B_{\alpha(a)}} \mu_b^{(a)} =: K_\mu > 0 \,, \text{ for all } a \in \ell \,, \tag{14}$$

(ii) the constant

$$K_a := \sup_{b \in B_{\alpha(a)}} \mu_b^{(a)} \, \mathbb{E}\big[T_{b \to p(b)}^{(a)}\big] \,, \tag{15}$$

satisfies

$$\frac{K_a^2}{\mathbb{E}\big[T_{a \to 0}^{(a)}\big]} \xrightarrow[a \to \infty]{} 0 \,. \tag{16}$$

Condition (i) ensures that transitions to the left remain non null asymptotically in a. It also provides a useful bound (see Lemma 4.1). The drift is given by (ii) and can be interpreted in two ways. Note first that by (13), K_a can be written

$$K_a = \sup_{b \in B_{\alpha(a)}} \frac{\pi^{(a)}(B_b)}{\pi^{(a)}(b)} \,. \tag{17}$$

Thus, condition (ii) can be seen as an exponential decay of the i.p.m. $\pi^{(a)}$ along the branch $B_{\alpha(a)}$ which contains a: we have $\pi^{(a)}(b) \le e^{-\gamma_a \, d(b)}$, with $\gamma_a = -\log(1 - 1/K_a)$, valid for all $b \in B_{\alpha(a)}$. That is for any state of any subbranch of the branch containing a, since by definition $B_{\alpha(a)} = \bigcup_{x \in \ell(a)} B_x$. In terms of energy profile, γ_a also provides a lower bound to the energy profile of any subbranch of the branch containing a. We refer to [5] for a more detailed discussion about these interpretations of our drift condition.

2.5   Results

For the fixed path ℓ which contains the sequence of a's going to infinity, the 0^ℓ-SD drift condition is a sufficient condition for the cut-off behavior of hitting times $T_{a \to 0}^{(a)}$:

Proposition 2.2. *If the family $X^{(a)}(t)$ of irreducible discrete time birth and death chains on I_a satisfies the 0^ℓ-SD condition, then $\lim\limits_{a \to \infty} \mathrm{Var}\big(T_{a \to 0}^{(a)} / \mathbb{E}\big[T_{a \to 0}^{(a)}\big]\big) = 0$.*

As a consequence, the random variables $T_{a \to 0}^{(a)}$ exhibit cut-off behavior at mean times.

Moreover, under this drift condition the typical time-scale $\mathbb{E}\big[T_{a \to 0}^{(a)}\big]$ of the abrupt convergence toward zero is negligible comparing to the mean times $\mathbb{E}\big[T_{0 \to a}^{(a)}\big]$ of the escape from zero to a:

Proposition 2.3. *If the family $X^{(a)}(t)$ of irreducible discrete time birth and death chains on I_a satisfies the 0^ℓ-SD condition, then $\lim\limits_{a \to \infty} \mathbb{E}\big[T_{a \to 0}^{(a)}\big] / \mathbb{E}\big[T_{0 \to a}^{(a)}\big] = 0$.*

This time-scales property is the key element of the presence of escape behavior for $T_{0 \to a}^{(a)}$. But, as seen in Section 3 of [5], this requires also that starting from any state of the state space, the mean hitting times of zero are negligible comparing to $\mathbb{E}\big[T_{0 \to a}^{(a)}\big]$, and that the sequence $T_{0 \to a}^{(a)} / \mathbb{E}\big[T_{0 \to a}^{(a)}\big]$ is uniformly integrable. The requirement that $\sup_{x \in I_a} \mathbb{E}\big[T_{x \to 0}^{(a)}\big] / \mathbb{E}\big[T_{0 \to a}^{(a)}\big] \xrightarrow[a \to \infty]{} 0$ is obtained if we suppose, in addition to the 0^ℓ-SD condition, that the mean hitting times of zero are comparable. That is if the ratio $\sup_{x \in I_a} \mathbb{E}\big[T_{x \to 0}^{(a)}\big] / \mathbb{E}\big[T_{a \to 0}^{(a)}\big]$ remains bounded for all a. For the uniform integrability of $T_{0 \to a}^{(a)} / \mathbb{E}\big[T_{0 \to a}^{(a)}\big]$, we need to control $\mathbb{E}\big[T_{0 \to a}^2\big]$ (see Corollary 3.5). A sufficient condition is to have $\inf_{b \in I_a} \mu_b^{(a)} > 0$ uniformly in a, and $K_a'^2 / \mathbb{E}\big[T_{0 \to a}^{(a)}\big] \xrightarrow[a \to \infty]{} 0$, where

$$K_a' := \sup_{b \in I_a} \mu_b^{(a)} \, \mathbb{E}\big[T_{b \to p(b)}^{(a)}\big] = \sup_{b \in I_a} \frac{\pi^{(a)}(B_b)}{\pi^{(a)}(b)} \,. \tag{18}$$

As above, the second condition is obtained if in addition to the 0^ℓ-SD condition, the ratio $K_a'^2 / \mathbb{E}\big[T_{a \to 0}^{(a)}\big]$ remains bounded for all a (see Section 4.3).

Theorem 2.4. *If the family $X^{(a)}(t)$ of irreducible discrete time birth and death chains on I_a satisfies the 0^ℓ-SD condition, and if there exist K, K' and K'_μ such that*

$$\frac{\sup_{x \in I_a} \mathbb{E}\big[T^{(a)}_{x \to 0}\big]}{\mathbb{E}\big[T^{(a)}_{a \to 0}\big]} \leq K \,, \quad \frac{K'^{\,2}_a}{\mathbb{E}\big[T^{(a)}_{a \to 0}\big]} \leq K' \,, \quad and \quad \inf_{b \in I_a} \mu^{(a)}_b =: K'_\mu > 0 \quad \forall a \in \ell \,, \qquad (19)$$

then:

1. *the random variables $T^{(a)}_{a \to 0}$ exhibit cut-off behavior at mean times.*

2. *the random variables $T^{(a)}_{0 \to a}$ exhibit escape-time behavior at mean times.*

Moreover, as the quantity K_a is a supremum taken over all sites of the branch $B_{\alpha(a)}$ which contains a, these assumptions are sufficient to have cut-off and escape behaviors between zero and every state b of $B_{\alpha(a)}$ for which $\mathbb{E}\big[T^{(a)}_{b \to 0}\big]$ diverges with $\mathbb{E}\big[T^{(a)}_{a \to 0}\big]$. Let us define

$$L(a) := \big\{\, b \in B_{\alpha(a)} : \exists\, C \text{ s.t. } \mathbb{E}\big[T^{(a)}_{b \to 0}\big] \geq C\, \mathbb{E}\big[T^{(a)}_{a \to 0}\big] \,\big\} \,. \qquad (20)$$

Corollary 2.5. *Theorem 2.4 holds for $T^{(a)}_{b \to 0}$ and $T^{(a)}_{0 \to b}$ for all $b \in L(a)$.*

As we will see in Section 4 (Lemmas 4.2 and 4.3), the conditions needed to obtain the two time scales $\big(\, \mathbb{E}\big[T^{(a)}_{a \to 0}\big]\, /\, \mathbb{E}\big[T^{(a)}_{0 \to a}\big] \xrightarrow[a \to \infty]{} 0 \,\big)$ – which are typical of escape behavior – are actually weaker than those which imply cut-off behavior. But unlike the case when the chain is defined on $\mathbb{N}$ where the drift condition implies both cut-off and escape behaviors (see [7], Theorem 2.2), here the 0^ℓ-SD condition implies only the cut-off behavior directly. As the chain started in zero could escape to other branches than $B_{\alpha(a)}$, we need in addition to have a "control" on that part of the trajectories from zero to a (conditions (19)) to get the escape behavior of $T^{(a)}_{0 \to a}$.

The assumption that $\sup_{x \in I_a} \mathbb{E}\big[T^{(a)}_{x \to 0}\big]/\mathbb{E}\big[T^{(a)}_{0 \to a}\big] \xrightarrow[a \to \infty]{} 0$ is the analogue of conditions (2.17) in Theorem 1 of [5] for chains on $\mathbb{Z}$. This is a central argument of the proof of Theorem 2 in [5] which gives sufficient conditions for escape behavior of $T^{(a)}_{0 \to a}$. The idea is that there exists an intermediate time scale Δ_a between $\mathbb{E}\big[T^{(a)}_{a \to 0}\big]$ and $\mathbb{E}\big[T^{(a)}_{0 \to a}\big]$ such that the walk will almost surely (in the limit $a \to \infty$) visit zero in an interval of length Δ_a. Thus, the walk started in zero will make a great number of unfructuous attempts to escape the attraction of zero before eventually hitting a. The exponential character (randomness without memory) of $T^{(a)}_{0 \to a}$ is a consequence of this situation: each trial is followed by a return to zero, and then by Markovianness, the walk starts again in the same situation.

Remark 2.6. If the tree I_a is reduced to the branch $B_{\alpha(a)}$ which contains a, the conditions of Theorem 2.4 reduce to 0^ℓ-SD plus $\sup_{x \in B_{\alpha(a)}} \mathbb{E}\big[T^{(a)}_{x \to 0}\big]/\mathbb{E}\big[T^{(a)}_{a \to 0}\big]$ bounded for all a.

Remark 2.7. Propositions 2.2 and 2.3 above also apply when the chains $X^{(a)}(t)$ are defined on an infinite tree I (i.e. when $I_a = I$ infinite for all a). We can also apply Theorem 2.4 in that case by considering the restrictions of the chains to $I_a := \{x \in I : d(x) \leq d(a)\}$.

Another feature of the complementarity between both phenomena, is that the successful final excursion which leads to a is "cut-off like" (see Theorem 1 and Section 7 in [5]): if we consider the hitting times $\widetilde{T}^{(a)}_{0 \to a}$ of a after the last visit to zero, and $\widetilde{T}^{(a)}_{a \to 0}$ (hitting time of zero after last visit to a), we can show by a reversibility argument that they have the same laws. Thus, cut-off and escape behaviors are related by time inversion.

2.6 Example: biased random walk on a regular tree

A regular tree is a tree I in which each node has the same number r of children. In its infinite version (also called Bethe lattice), the degree (number of neighbors) of each node is $r + 1$, except the root which is of degree r. When we consider the finite version $I_a := \{x \in I : d(x) \leq d(a)\}$ for some a, the boundary nodes (sites x s.t. $d(x) = d(a)$) – also called leaves of the tree – are dead ends (of degree 1). In that case, we talk about a Cayley tree. We refer to [9, 10] for discussions about the differences between Bethe lattice and Cayley tree, and to Chapter 5 of [4] for the biased random walk on a regular tree.

We consider biased random walks $X^{(a)}(t)$ on I_a which transitions are defined by

$$lx := \frac{1}{\lambda + r} \quad \text{and} \quad \mu_x := \frac{\lambda}{\lambda + r} \,, \tag{21}$$

with $0 < \lambda < \infty$, for each site $x \neq 0$. As transitions must sum to one in each site, we add also "waiting probabilities" at root and leaves: $\kappa_0 := \lambda/(\lambda + r)$ and $\kappa_x := r/(\lambda + r)$ when $d(x) = d(a)$. The parameter λ represents the bias of the random walk: at each step, the chain moves toward the root of the tree with probability $\lambda/(\lambda+r)$ and in the opposite direction with probability $r/(\lambda+r)$. Thus, there is a concentration phenomenon near the root level when $\lambda > r$ (localization), and near the leaves level when $\lambda < r$ (delocalization).

Due to the regular structure of the tree, we can compute the quantities involved in the 0^ℓ-SD condition, and determine when it is satisfied or not. Let us start with the invariant measure π. By (10), for all x we have

$$\pi(x) = \pi(0) \prod_{y \in \ell(x)} \frac{ly}{\mu_y} = \pi(0) \prod_{y \in \ell(x)} \frac{1}{\lambda + r} \frac{\lambda + r}{\lambda} = \pi(0) \frac{1}{\lambda^{d(x)}} \,, \tag{22}$$

with $\pi(0) = 1 / \sum_{x \in I_a} \frac{1}{\lambda^{d(x)}}$.

Remark 2.8. The computations of mean return times using Kac formula $\left(\mathbb{E}[T_x] = 1/\pi(x) \right.$, where $T_x := \inf\{t > 0 : X^x(t) = x\}$ is the first return time to $x \left. \right)$ illustrates the localization phenomenon for the biased random walk on a regular tree:

- localization is related to positive recurrence of the walk when it takes place on an infinite tree. Lyons showed that the biased walk is recurrent (return times to any state almost surely finite) when the bias λ exceeds the branching number of the tree (here $\lambda > r$) and transient when $\lambda < r$ (see [11]).
  Moreover, when $\lambda > r$, the chain is positive recurrent (finite mean return times to any state). Using the Kac formula and since for each k there are r^k sites y in I such that $d(y) = k$, we get

$$\mathbb{E}[T_x] = \lambda^{d(x)} \sum_{y \in I} \frac{1}{\lambda^{d(y)}} = \lambda^{d(x)} \sum_{k=0}^{\infty} \left(\frac{r}{\lambda} \right)^k < \infty, \quad \text{if } \lambda > r. \tag{23}$$

- in the finite case, although the walk on I_a is always positive recurrent, the localization phenomenon is characterized by very different asymptotical behaviors of mean return times to zero and a respectively. When $\lambda > r$ the mean return time to zero remains bounded as a goes to infinity (by Kac formula, we get $\mathbb{E}[T_0^{(a)}] \sim \lambda/(\lambda - r)$ for large a). In opposition, the mean return time to a diverges with a $(\mathbb{E}[T_a^{(a)}] \sim \lambda^{d(a)+1}/(\lambda - r)$ for large a).

Let us now compute the key quantity $\frac{\pi(B_x)}{\pi(x)}$ which appears in both expressions of K_a and $\mathbb{E}\big[T_{a\to 0}^{(a)}\big]$ (see (17), and Proposition 3.3 below). We write

$$\frac{\pi(B_x)}{\pi(x)} = \sum_{b\in B_x} \frac{\pi(b)}{\pi(x)} = \sum_{b\in B_x} \frac{\lambda^{d(x)}}{\lambda^{d(b)}} = \sum_{k=0}^{d(a)-d(x)} \left(\frac{r}{\lambda}\right)^k, \tag{24}$$

since there are r^k sites b in B_x such that $d(b) = d(x) + k$.

When $\lambda \neq r$, this is equal to $\frac{\lambda}{\lambda - r}\left(1 - \left(\frac{r}{\lambda}\right)^{d(a)-d(x)+1}\right)$, and thus

$$K_a := \sup_{x\in\ell(a)} \sup_{b\in B_x} \frac{\pi(B_b)}{\pi(b)} = \frac{\lambda}{\lambda - r}\left(1 - \left(\frac{r}{\lambda}\right)^{d(a)}\right), \tag{25}$$

and

$$\begin{aligned}
\mathbb{E}\big[T_{a\to 0}^{(a)}\big] &= \sum_{x\in\ell(a)} \frac{\pi(B_x)}{\mu_x\,\pi(x)} = \frac{\lambda+r}{\lambda-r} \sum_{x\in\ell(a)} \left(1 - \left(\frac{r}{\lambda}\right)^{d(a)-d(x)+1}\right) \\
&= \frac{\lambda+r}{\lambda-r}\left(d(a) - \left(\frac{r}{\lambda}\right)^{d(a)} \sum_{k=0}^{d(a)-1} \left(\frac{\lambda}{r}\right)^k\right) \\
&= \frac{\lambda+r}{\lambda-r}\left(d(a) + \frac{r}{\lambda-r}\left(\left(\frac{r}{\lambda}\right)^{d(a)} - 1\right)\right).
\end{aligned} \tag{26}$$

In the localization case ($\lambda > r$), we have $\mathbb{E}\big[T_{a\to 0}^{(a)}\big] \sim \frac{\lambda+r}{\lambda-r}\,d(a)$, and $K_a \to K := \frac{\lambda}{\lambda-r}$ as $a \to \infty$. The 0^ℓ-SD condition is then satisfied, and we have cut-off and escape behavior between the root and any leave of I_a, since for symmetry reasons, the above computations are the same for any state of I_a. This result leads to a better understanding of what happens in that case: since $T_{a\to 0}^{(a)}$ exhibits cut-off behavior at mean times $\mathbb{E}\big[T_{a\to 0}^{(a)}\big] \sim d(a)$, the walk started in a will go in an almost deterministic way to zero before eventually returning to a. Moreover, the typical time scale of cut-off trajectories is negligible comparing to that of escape trajectories (computations give $\mathbb{E}\big[T_{0\to a}^{(a)}\big] \sim \lambda^{d(a)}$ for large a, which is of the same order than $\mathbb{E}\big[T_a^{(a)}\big]$).

In the delocalization case ($\lambda < r$), K_a and $\mathbb{E}\big[T_{a\to 0}^{(a)}\big]$ are both exponential of order $(r/\lambda)^{d(a)}$ for large a, and then there is no strong drift. The 0^ℓ-SD condition is also not satisfied when $\lambda = r$. Computations in that case show that $K_a = d(a)$ and $\mathbb{E}\big[T_{a\to 0}^{(a)}\big] = \big((\lambda+r)/2\lambda\big)\,d(a)\big(d(a)+1\big)$ and thus the ratio $K_a^2/\mathbb{E}\big[T_{a\to 0}^{(a)}\big]$ does not converge to zero when a goes to ∞.

3 Mean hitting times

We already mentioned (without proving it) in Section 2.3, that the expression of $\mathbb{E}\big[T_{x\to p(x)}\big]$ in terms of the invariant measure π (equation (13)) is a key ingredient in our computations. However, we will also need exact expressions for the first and second moments of hitting times between sites along the path from zero to some fixed state $a \in J$. We give a list of these expressions in Section 3.1. We start with first and second moments for $T_{x\to p(x)}$ (Lemma 3.1), and $T_{p(x)\to x}$ (Lemma 3.2). Proposition 3.3 gives the mean hitting times between sites $a_i \in \ell(a)$ for some $a \in J$, and Proposition 3.4 their second moments. Finally we give an idea of the proofs in Section 3.2.

3.1 Exact expressions

Lemma 3.1. *For any $x \in J$*

$$\mathbb{E}\left[T_{x \to p(x)}\right] = \frac{\pi(B_x)}{\mu_x \pi(x)} , \tag{27}$$

$$\mathbb{E}\left[T_{x \to p(x)}^2\right] = \frac{2}{\mu_x \pi(x)} \sum_{b \in B_x} \frac{\pi(B_b)^2}{\mu_b \pi(b)} - \mathbb{E}\left[T_{x \to p(x)}\right] . \tag{28}$$

Lemma 3.2. *For any $x \in J$*

$$\mathbb{E}\left[T_{p(x) \to x}\right] = \frac{1 - \pi(B_x)}{\mu_x \pi(x)} , \tag{29}$$

$$\mathbb{E}\left[T_{p(x) \to x}^2\right] = \frac{2}{\mu_x \pi(x)} \left(\sum_{c \in C_x} \frac{\pi(B_c)^2}{\mu_c \pi(c)} + \sum_{b \in \ell(x)} \frac{\pi\left(\overline{B}_b\right)^2}{\mu_b \pi(b)} \right) - \mathbb{E}\left[T_{p(x) \to x}\right] . \tag{30}$$

Proposition 3.3. *For $a \in J$, recall that we denote by a_i the i-th parent of a $\left(a_i = p^i(a)\right)$.*
Let $0 \leq j < n \leq d(a)$, we have

$$\mathbb{E}\left[T_{a_j \to a_n}\right] = \sum_{k=j}^{n-1} \frac{\pi(B_{a_k})}{\mu_{a_k} \pi(a_k)} , \tag{31}$$

$$\mathbb{E}\left[T_{a_n \to a_j}\right] = \sum_{k=j}^{n-1} \frac{1 - \pi(B_{a_k})}{\mu_{a_k} \pi(a_k)} . \tag{32}$$

And thus

$$\mathbb{E}\left[T_{a_j \to a_n}\right] + \mathbb{E}\left[T_{a_n \to a_j}\right] = \sum_{k=j}^{n-1} \frac{1}{\mu_{a_k} \pi(a_k)} . \tag{33}$$

Proposition 3.4. *For $0 \leq j < n \leq d(a)$, we have*

$$\mathbb{E}\left[T_{a_j \to a_n}^2\right] = \sum_{k=j}^{n-1} \frac{2}{\mu_{a_k} \pi(a_k)} \left(\sum_{b \in B_{a_k}} \frac{\pi(B_b)^2}{\mu_b \pi(b)} + \pi(B_{a_k}) \mathbb{E}\left[T_{a_{k+1} \to a_n}\right] \right) - \mathbb{E}\left[T_{a_j \to a_n}\right] , \tag{34}$$

and

$$\mathbb{E}\left[T_{a_n \to a_j}^2\right] = \sum_{k=j}^{n-1} \frac{2}{\mu_{a_k} \pi(a_k)} \left(\sum_{c \in C_{a_k}} \frac{\pi(B_c)^2}{\mu_c \pi(c)} + \sum_{b \in \ell(a_k)} \frac{\pi\left(\overline{B}_b\right)^2}{\mu_b \pi(b)} + \pi\left(\overline{B}_{a_k}\right) \mathbb{E}\left[T_{a_k \to a_j}\right] \right) - \mathbb{E}\left[T_{a_n \to a_j}\right] . \tag{35}$$

Corollary 3.5. *For $0 \leq j < n \leq d(a)$, we have*

$$\mathbb{E}\left[T_{a_n \to a_j}^2\right] \leq \mathbb{E}\left[T_{a_n \to a_j}\right] \left(2 \left(\frac{K_a'^2}{K_\mu'} + \mathbb{E}\left[T_{0 \to a_j}\right] \right) - 1 \right) . \tag{36}$$

Proof
This a consequence of equation (35). We first use the fact that $\overline{B}_b \subseteq \overline{B}_{a_k}$ for all $b \in \ell(a_k)$ to write

$$\sum_{b \in \ell(a_k)} \frac{\pi\left(\overline{B}_b\right)^2}{\mu_b \pi(b)} \leq \pi\left(\overline{B}_{a_k}\right) \sum_{b \in \ell(a_k)} \frac{\pi\left(\overline{B}_b\right)}{\mu_b \pi(b)} = \pi\left(\overline{B}_{a_k}\right) \mathbb{E}\left[T_{0 \to a_k}\right] . \tag{37}$$

Then, as $C_{a_k} \subset \overline{B}_{a_k}$, and since by definition $K'_a = \sup_{b \in I_a} \frac{\pi(B_b)}{\pi(b)}$ and $K'_\mu := \inf_{b \in I_a} \mu_b^{(a)}$, we have

$$\sum_{c \in C_{a_k}} \frac{\pi(B_c)^2}{\mu_c \, \pi(c)} \leq \frac{K_a'^2}{K'_\mu} \sum_{c \in C_{a_k}} \pi(c) = \frac{K_a'^2}{K'_\mu} \pi(C_{a_k}) \leq \frac{K_a'^2}{K'_\mu} \pi(\overline{B}_{a_k}) . \tag{38}$$

Finally, from equation (35)

$$\mathbb{E}[T^2_{a_n \to a_j}] \leq 2 \sum_{k=j}^{n-1} \frac{\pi(\overline{B}_{a_k})}{\mu_{a_k} \pi(a_k)} \left(\frac{K_a'^2}{K'_\mu} + \mathbb{E}[T_{0 \to a_k}] + \mathbb{E}[T_{a_k \to a_j}] \right) - \mathbb{E}[T_{a_n \to a_j}] . \tag{39}$$

Hence the result, since $T_{0 \to a_k} + T_{a_k \to a_j} = T_{0 \to a_j}$ and then using equation (32).

$\square$

3.2 Proofs

The above formulas can be proven via the standard method of difference equations which rely on the following decomposition of $\mathbb{E}[F(T_{a_k \to a_n})]$ for any function F:

Lemma 3.6. *For all $0 < k < d(a)$, $k \neq n$, we have*

$$\mathbb{E}[F(T_{a_k \to a_n})] = \mu_{a_k} \mathbb{E}[F(T_{a_{k+1} \to a_n} + 1)] + \sum_{c \in \sigma(a_k)} lc \, \mathbb{E}[F(T_{c \to a_n} + 1)] . \tag{40}$$

Proof

This is obtained by decomposing according to the first step of the chain

$$\mathbb{E}[F(T_{a_k \to a_n})] = \mu_{a_k} \mathbb{E}[F(T_{a_k \to a_n}) \mid X(1) = a_{k+1}] + \sum_{c \in \sigma(a_k)} lc \, \mathbb{E}[F(T_{a_k \to a_n}) \mid X(1) = c] , \tag{41}$$

and by Markovianness, for $x = a_{k+1}$ and $c \in \sigma(a_k)$, we have

$$
\begin{aligned}
\mathbb{E}[F(T_{a_k \to a_n}) \mid X(1) = x] &= \sum_{l \geq 1} F(l) \, \mathbb{P}(T_{a_k \to a_n} = l \mid X(1) = x) \\
&= \sum_{l \geq 1} F(l) \, \mathbb{P}(T_{x \to a_n} = l - 1) \\
&= \sum_{l \geq 0} F(l+1) \, \mathbb{P}(T_{x \to a_n} = l) \\
&= \mathbb{E}[F(T_{x \to a_n} + 1)] .
\end{aligned}
\tag{42}
$$

Hence the result.

$\square$

Together with Lemma 3.1 which is the discrete analogue of Lemma 1 of [13] for continuous chains, Lemma 3.6 allows us to prove Propositions 3.3 and 3.4: we fix $0 \leq n \leq d(a)$ and apply equation (40) with $F(x) = x$ (resp. $F(x) = x^2$) to obtain recursive equations for $\mathbb{E}[T_{a_k \to a_n}]$, (resp. $\mathbb{E}[T^2_{a_k \to a_n}]$) with $0 < k < d(a)$. Then we can iterate these equations in both directions up to boundary conditions – which are particular cases of (40) for $k = 0$ and $d(a)$ – and solve them using reversibility and equations (27) and (28).

Another way of proving both propositions is to use, together with Lemmas 3.1 and 3.2, the fact that $T_{a_j \to a_n}$ (resp. $T_{a_n \to a_j}$) is the sum of independant $T_{x \to p(x)}$ (resp. $T_{p(x) \to x}$):

$$T_{a_j \to a_n} = \sum_{k=j}^{n-1} T_{a_k \to p(a_k)} \quad \text{and} \quad T_{a_n \to a_j} = \sum_{k=j}^{n-1} T_{p(a_k) \to a_k} . \tag{43}$$

We give now a sketch of those proofs.

First moments:
Althought they can be proven directly by difference equations, formulas (27) and (29) are proven in [12] (Proposition 9.8) by an alternative method. And thus (31) and (32) of Proposition 3.3 follow from (43).

Second moments:
The proof of (28) follows the same lines as the ones of Lemma 1 in [13], where the application of equation (40) gives the recursive equation $\mu_x\, t_x = 1 + \sum_{c\in\sigma(x)} lc\, t_c$ for $t_x = \mathbb{E}\big[T_{x\to p(x)}\big]$. We now apply (40) with $F(x) = x^2$, $a_k = x$ and $a_n = p(x)$ to obtain the recursive equation for $\mathbb{E}\big[T^2_{x\to p(x)}\big]$:

$$\mu_x\, \mathbb{E}\big[T^2_{x\to p(x)}\big] \;=\; 2\,\mu_x(\mathbb{E}\big[T_{x\to p(x)}\big])^2 \;+\; \sum_{c\in\sigma(x)} lc\,\mathbb{E}\big[T^2_{c\to p(c)}\big] \;-\; 1\,. \tag{44}$$

Then we prove by induction that formula (28) satisfies the above equation. As in [13], we have first to suppose that the tree is finite, since the first step of induction is to consider a leave x (that is when $\sigma(x) = \varnothing$). For the infinite case, we consider the limit $n\to\infty$ of restrictions to the truncated trees $I_n := \{x\in I : d(x)\le n\}$ of level n.
For equation (30), we prove by induction that it satisfies the recursive equation obtained from (40) with $F(x) = x^2$, $a_k = p(x)$ and $a_n = x$. We don't need to restrict ourselves to finite trees in that case, since the first step of induction is to consider x such that $p(x) = 0$.
Finally, (34) and (35) of Proposition 3.4 follow from (43).

4   Proofs of the main results

The quantities involved in this section are defined with respect to the chain $X^{(a)}(t)$ on I_a, but for the sake of notations we will omit the upper index (a). We will prove Propositions 2.2 and 2.3 through Lemmas 4.2 and 4.3 which are a little bit stronger. They involve the quantities

$$Q(x) := \frac{1}{\pi\,(B_x)} \sum_{b\in B_x} \frac{\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} \quad ; \quad Q_a := \sup_{x\in\ell(a)} Q(x)\,, \tag{45}$$

and

$$R_a := \sum_{b\in\ell(a)} \frac{\pi\,(B_b)^2}{\mu_b\,\pi\,(b)}\,, \tag{46}$$

which for all a satisfy the inequalities

Lemma 4.1.
$$R_a \;\le\; Q_a \;\le\; K_a^2\,/\,K_\mu\,. \tag{47}$$

Proof of Lemma 4.1
As $\ell(a) \subset B_{a_{d(a)-1}}$, we have

$$R_a \;\le\; \sum_{b\in B_{a_{d(a)-1}}} \frac{\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} \;=\; \pi\big(B_{a_{d(a)-1}}\big)\,Q(a_{d(a)-1})\,, \tag{48}$$

and thus $R_a \le Q_a$, since $\pi\big(B_{a_{d(a)-1}}\big)\,Q(a_{d(a)-1}) \le Q(a_{d(a)-1}) \le Q_a$.
For the right-hand side of (47), we write

$$\frac{1}{\pi\,(B_x)} \sum_{b\in B_x} \frac{\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} \;\le\; \frac{1}{\pi\,(B_x)}\,\frac{K(x)^2}{K_\mu} \sum_{b\in B_x} \pi\,(b) \;=\; \frac{K(x)^2}{K_\mu}\,, \tag{49}$$

where $K(x) = \sup_{b\in B_x}\left(\pi\,(B_b)\,/\,\pi\,(b)\right)$.
We conclude by taking the supremum over $x\in\ell(a)$, since by (17), $K_a = \sup_{x\in\ell(a)} K(x)$.

$\square$

4.1 Proof of Proposition 2.2

We recall that

$$\lim_{a\to\infty} \text{Var}\Big(\frac{T_{a\to 0}}{\mathbb{E}\big[T_{a\to 0}\big]}\Big) = 0 \tag{50}$$

is a sufficient condition to prove the cut-off behavior of the variables $T_{a\to 0}$ (see [5], Section 3). Since $Q_a \le K_a^2 / K_\mu$ (see Lemma 4.1 above), and since by the 0^ℓ-SD hypothesis the ratio $K_a^2/\mathbb{E}\big[T_{a\to 0}\big]$ tends to zero, the following result implies Proposition 2.2.

Lemma 4.2. *If*

$$\frac{Q_a}{\mathbb{E}\big[T_{a\to 0}\big]} \xrightarrow[a\to\infty]{} 0 \, , \tag{51}$$

then $\text{Var}\big(T_{a\to 0}/\mathbb{E}\big[T_{a\to 0}\big]\big) \xrightarrow[a\to\infty]{} 0 \, .$

Proof of Lemma 4.2
For all $a \in J$, we have

$$
\begin{aligned}
\text{Var}\Big(T_{a\to 0}\Big) &= \sum_{x\in\ell(a)} \text{Var}\Big(T_{x\to p(x)}\Big) \\
&\le \sum_{x\in\ell(a)} \mathbb{E}\big[T_{x\to p(x)}^2\big] \\
&\le \sum_{x\in\ell(a)} \frac{2}{\mu_x \, \pi\,(x)} \sum_{b\in B_x} \frac{\pi\,(B_b)^2}{\mu_b \, \pi\,(b)} \, ,
\end{aligned}
\tag{52}
$$

where the second inequality comes from equation (28) of Lemma 3.1.
Thus

$$\text{Var}\Big(T_{a\to 0}\Big) \le \sum_{x\in\ell(a)} \frac{2\,\pi\,(B_x)}{\mu_x\,\pi\,(x)} \, Q(x) \le 2\,Q_a \sum_{x\in\ell(a)} \frac{\pi\,(B_x)}{\mu_x\,\pi\,(x)} \, , \tag{53}$$

and by (27)

$$\text{Var}\Big(\frac{T_{a\to 0}}{\mathbb{E}\big[T_{a\to 0}\big]}\Big) \le \frac{2\,Q_a}{\mathbb{E}\big[T_{a\to 0}\big]} \xrightarrow[a\to\infty]{} 0 \, . \tag{54}$$

$\square$

4.2 Proof of Proposition 2.3

As above, by Lemma 4.1 ($R_a \le K_a^2 / K_\mu$) together with the 0^ℓ-SD condition, the following lemma implies Proposition 2.3.

Lemma 4.3. *If*

$$\frac{R_a}{\mathbb{E}\big[T_{a\to 0}\big]} \xrightarrow[a\to\infty]{} 0 \, , \tag{55}$$

then $\mathbb{E}\big[T_{a\to 0}\big] / \mathbb{E}\big[T_{0\to a}\big] \xrightarrow[a\to\infty]{} 0 \, .$

Proof of Lemma 4.3
It is equivalent to show that

$$\Gamma_a \; := \; \frac{\mathbb{E}\big[T_{a\to 0}\big]}{\mathbb{E}\big[T_{a\to 0}\big] + \mathbb{E}\big[T_{0\to a}\big]} \xrightarrow[a\to\infty]{} 0 \, . \tag{56}$$

12

By (31) and (33), we have

$$\Gamma_a = \frac{\displaystyle\sum_{x \in \ell(a)} \frac{\pi\left(B_x\right)}{\mu_x\,\pi\left(x\right)}}{\displaystyle\sum_{x \in \ell(a)} \frac{1}{\mu_x\,\pi\left(x\right)}} \quad. \tag{57}$$

We now consider the probability measure on $\ell(a)$ defined by expectations

$$\mathcal{E}(F) := \frac{\displaystyle\sum_{x \in \ell(a)} \frac{F(x)}{\mu_x\,\pi\left(x\right)}}{\displaystyle\sum_{x \in \ell(a)} \frac{1}{\mu_x\,\pi\left(x\right)}} \quad, \tag{58}$$

and the inequality $\mathcal{E}\left(F\right)^2 \le \mathcal{E}\left(F^2\right)$ for $F(x) = \pi\left(B_x\right)$ implies that $\Gamma_a \le R_a \,/\, \mathbb{E}\left[T_{a\to 0}\right]$, which concludes the proof.

$\square$

4.3 Proof of Theorem 2.4

The cut-off behavior of $T_{a\to 0}$ was already proven in Proposition 2.2. For the escape behavior of $T_{0\to a}$, we need to prove that $\sup_{x \in I_a} \mathbb{E}\left[T_{x\to 0}\right]/\mathbb{E}\left[T_{0\to a}\right]$ tends to zero and that the sequence $T_{0\to a}\,/\,\mathbb{E}\left[T_{0\to a}\right]$ is uniformly integrable (see [5], Theorem 2).
For the first point, we have

$$\frac{\sup_{x \in I_a} \mathbb{E}\left[T_{x\to 0}\right]}{\mathbb{E}\left[T_{0\to a}\right]} = \frac{\sup_{x \in I_a} \mathbb{E}\left[T_{x\to 0}\right]}{\mathbb{E}\left[T_{a\to 0}\right]} \frac{\mathbb{E}\left[T_{a\to 0}\right]}{\mathbb{E}\left[T_{0\to a}\right]} \le K \frac{\mathbb{E}\left[T_{a\to 0}\right]}{\mathbb{E}\left[T_{0\to a}\right]} \xrightarrow[a\to\infty]{} 0\,, \tag{59}$$

as a consequence of (19) and Proposition 2.3.
For the uniform integrability, we apply Corollary 3.5 with $n = d(a)$ and $j = 0$ to get

$$\frac{\mathbb{E}\left[T_{0\to a}^2\right]}{\mathbb{E}\left[T_{0\to a}\right]^2} \le 2 + \frac{2}{K_\mu'} \frac{K_a'^2}{\mathbb{E}\left[T_{0\to a}\right]} \le 2 + o(1)\,. \tag{60}$$

The second inequality is due to (19) and Proposition 2.3, and thus the sequence $T_{0\to a}\,/\,\mathbb{E}\left[T_{0\to a}\right]$ is uniformly integrable.

4.4 Proof of Corollary 2.5

The key element of this result, is that K_b is equal to K_a, as $B_{\alpha(b)} = B_{\alpha(a)}$ for $b \in L(a)$. Then if the 0^ℓ-SD condition holds for a, it also holds for every state b of $L(a)$:

$$\frac{K_b^2}{\mathbb{E}\left[T_{b\to 0}\right]} = \frac{K_a^2}{\mathbb{E}\left[T_{b\to 0}\right]} \le \frac{K_a^2}{C\,\mathbb{E}\left[T_{a\to 0}\right]} \xrightarrow[a\to\infty]{} 0\,. \tag{61}$$

Since equation (47) is valid for any state, we can apply Lemma 4.2 and Proposition 2.2 holds for all b: that is $T_{b\to 0}$ exhibits cut-off behavior for every $b \in L(a)$.
It remains to prove the escape behavior of $T_{0\to b}$. Let $b \in L(a)$, using (19) we have

$$\frac{\sup_{x \in I_a} \mathbb{E}\left[T_{x\to 0}\right]}{\mathbb{E}\left[T_{0\to b}\right]} = \frac{\sup_{x \in I_a} \mathbb{E}\left[T_{x\to 0}\right]}{\mathbb{E}\left[T_{a\to 0}\right]} \frac{\mathbb{E}\left[T_{a\to 0}\right]}{\mathbb{E}\left[T_{b\to 0}\right]} \frac{\mathbb{E}\left[T_{b\to 0}\right]}{\mathbb{E}\left[T_{0\to b}\right]} \le \frac{K}{C} \frac{\mathbb{E}\left[T_{b\to 0}\right]}{\mathbb{E}\left[T_{0\to b}\right]}\,, \tag{62}$$

and this tends to zero since by Lemma 4.3, Proposition 2.3 holds for b.

Now by Corollary 3.5, we have

$$\frac{\mathbb{E}\big[T_{0\to b}^2\big]}{\mathbb{E}\big[T_{0\to b}\big]^2} \leq 2 + \frac{2}{K_\mu}\, \frac{{K'_a}^2}{\mathbb{E}\big[T_{0\to b}\big]} \leq 2 + o(1)\,, \tag{63}$$

since by (19) and Proposition 2.3

$$\frac{{K'_a}^2}{\mathbb{E}\big[T_{0\to b}\big]} = \frac{{K'_a}^2}{\mathbb{E}\big[T_{a\to \circ}\big]}\, \frac{\mathbb{E}\big[T_{a\to \circ}\big]}{\mathbb{E}\big[T_{b\to \circ}\big]}\, \frac{\mathbb{E}\big[T_{b\to \circ}\big]}{\mathbb{E}\big[T_{0\to b}\big]} \leq \frac{K'}{C}\, \frac{\mathbb{E}\big[T_{b\to \circ}\big]}{\mathbb{E}\big[T_{0\to b}\big]} \xrightarrow[a\to\infty]{} 0\,. \tag{64}$$

Thus $T_{0\to b}\,/\,\mathbb{E}\big[T_{0\to b}\big]$ is uniformly square integrable, and $T_{0\to b}$ exhibits escape behavior.

References

[1] Aldous, D.: Random walks on finite groups and rapidly mixing Markov chains. In: Séminaire de probabilités XVII, *Lecture Notes in Math.* vol. 986, pp. 243–297. Springer, Berlin (1983)

[2] Aldous, D., Diaconis, P.: Shuffling cards and stopping times. Amer. Math. Monthly **93**(5), 333–348 (1986)

[3] Aldous, D., Diaconis, P.: Strong uniform times and finite random walks. Adv. in Appl. Math. **8**(1), 69–97 (1987)

[4] Aldous, D., Fill J.: Reversible Markov Chains and Random Walks on Graphs. Monograph in preparation *http://www.stat.berkeley.edu/users/aldous/RWG/book.html*

[5] Barrera, J., Bertoncini, O., Fernández, R.: Abrupt Convergence and Escape Behavior for Birth and Death Chains J. Stat. Phys. **137**, 595-623 (2009)

[6] Bertoncini, O.: Convergence abrupte et métastabilité. Ph.D. thesis, Université de Rouen (2007)

[7] Bertoncini, O., Barrera, J., Fernández, R.: Cut-off and exit from metastability: two sides of the same coin. C. R. Acad. Sci. Paris, Ser. I **346**, 691–696 (2008)

[8] Cassandro, M., Galves, A., Olivieri, E., Vares, M.E.: Metastable behavior of stochastic dynamics: a pathwise approach. J. Statist. Phys. **35**(5-6), 603–634 (1984)

[9] Dorogovtsev, S.N., Goltsev, A. V., Mendes, J.F.F.: Critical phenomena in complex networks. Rev. Mod. Phys., 80, 1275 (2008)

[10] Dorogovtsev, S.N.: Lectures on Complex Networks. Oxford University Press, Oxford (2010)

[11] Lyons, R.: Random Walks and Percolation on Trees. Ann. Probab. **18**(3), 931–958 (1990)

[12] Woess, W.: Denumerable Markov Chains. Generating functions, boundary theory, random walks on trees. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich (2009)

[13] Martínez, S., Ycart, B.: Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space. Adv. in Appl. Probab. **33**(1), 188–205 (2001)

Appendix: Difference equations for mean hitting times

We give here the detailed proofs of Propositions 3.3 and 3.4 of Section 3 using the standart method of difference equations.

Proof of Proposition 3.3
Let $a \in J$ and $0 \leq n \leq d(a)$ be fixed, and define $D_x := \mathbb{E}\big[F(T_{x \to a_n})\big]$ for all $x \in \ell(a)$. Since by (9) the transitions out of each site sum to 1, equation (40), with $F(T) = T$, can be rewritten

$$D_{a_k} = \mu_{a_k} D_{a_{k+1}} + \lambda_{a_{k-1}} D_{a_{k-1}} + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc\, D_c + 1 \,, \tag{65}$$

valid for all $0 < k < d(a)$, $k \neq n$, with the boundary conditions

$$D_{a_0} := \mathbb{E}\big[T_{a \to a_n}\big] = \mu_a D_{a_1} + \sum_{c \in \sigma(a)} lc\, D_c + 1 \,, \tag{66}$$

$$D_{a_{d(a)}} := \mathbb{E}\big[T_{0 \to a_n}\big] = \lambda_{a_{d(a)-1}} D_{a_{d(a)-1}} + \sum_{\substack{c \in \sigma(0) \\ c \neq a_{d(a)-1}}} lc\, D_c + 1 \,, \tag{67}$$

and

$$D_{a_n} := \mathbb{E}\big[T_{a_n \to a_n}\big] = 0 \,. \tag{68}$$

Now, for $c \in \sigma(a_k), c \neq a_{k-1}$, the hitting time $T_{c \to a_n}$ of a_n starting from c is the sum of the time $t_c := T_{c \to p(c)}$ from c to $p(c) = a_k$ and $T_{a_k \to a_n}$, i.e. $D_c = t_c + D_{a_k}$.
Using (9) this leads to the following difference equation

$$\big(\mu_{a_k} + \lambda_{a_{k-1}}\big) D_{a_k} = \mu_{a_k} D_{a_{k+1}} + \lambda_{a_{k-1}} D_{a_{k-1}} + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc\, t_c + 1 \,, \tag{69}$$

with boundary conditions (68) and

$$\mu_a D_{a_0} = \mu_a D_{a_1} + \sum_{c \in \sigma(a)} lc\, t_c + 1 \,, \tag{70}$$

$$\lambda_{a_{d(a)-1}} D_{a_{d(a)}} = \lambda_{a_{d(a)-1}} D_{a_{d(a)-1}} + \sum_{\substack{c \in \sigma(0) \\ c \neq a_{d(a)-1}}} lc\, t_c + 1 \,, \tag{71}$$

Proof of (31): Suppose that $0 \leq j < n$.
The idea is to write D_{a_j} as

$$D_{a_j} = D_{a_0} - \sum_{k=0}^{j-1} \big(D_{a_k} - D_{a_{k+1}}\big) \,. \tag{72}$$

Together with boundary condition (68) which gives $D_{a_0} = \sum_{k=0}^{n-1} \big(D_{a_k} - D_{a_{k+1}}\big)$, this leads to

$$D_{a_j} = \sum_{k=j}^{n-1} \big(D_{a_k} - D_{a_{k+1}}\big) \,. \tag{73}$$

Let $\alpha_k := \dfrac{1}{\mu_{a_k}} \Big(1 + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc\, t_c\Big)$ and $\underline{k} := \dfrac{\lambda_{a_{k-1}}}{\mu_{a_k}}$ for $k = 1, \ldots, n-1$.
We can rewrite (69) in the form

$$D_{a_k} - D_{a_{k+1}} = \underline{k}\, \big(D_{a_{k-1}} - D_{a_k}\big) + \alpha_k \,, \tag{74}$$

and then iterate this difference equation down to $k = 0$ (recall that $a_0 = a$), to obtain

$$D_{a_k} - D_{a_{k+1}} = \prod_{i=1}^{k} \underline{i}(D_{a_0} - D_{a_1}) + \sum_{l=1}^{k} \alpha_l \prod_{i=l+1}^{k} \underline{i}. \tag{75}$$

By condition (70), we have

$$D_{a_0} - D_{a_1} = \frac{1}{\mu_a} \left(1 + \sum_{c \in \sigma(a)} lc\, t_c \right) =: \alpha_0\,, \tag{76}$$

Thus

$$D_{a_k} - D_{a_{k+1}} = \sum_{l=0}^{k} \alpha_l \prod_{i=l+1}^{k} \underline{i}. \tag{77}$$

And finally

$$D_{a_j} = \sum_{k=j}^{n-1} \left(D_{a_k} - D_{a_{k+1}} \right) = \sum_{k=j}^{n-1} \sum_{l=0}^{k} \alpha_l \prod_{i=l+1}^{k} \underline{i}. \tag{78}$$

The quantities $\alpha_l \prod_{i=l+1}^{k} \underline{i}$ can be expressed in terms of the invariant measure π: by definition (10) of the i.p.m. π, we have

$$\frac{1}{\mu_{a_l}} \prod_{i=l+1}^{k} \frac{\lambda_{a_{i-1}}}{\mu_{a_i}} = \frac{1}{\mu_{a_k}} \prod_{i=l}^{k-1} \frac{\lambda_{a_i}}{\mu_{a_i}} = \frac{1}{\mu_{a_k}} \frac{\pi(a_l)}{\pi(a_k)}\,, \tag{79}$$

and using reversibility and expression (13) for $t_c := \mathbb{E}\left[T_{c \to p(c)}\right]$, we have

$$lc\, t_c = \frac{lc\, \pi(B_c)}{\mu_c\, \pi(c)} = \frac{\pi(B_c)}{\pi(p(c))} = \frac{\pi(B_c)}{\pi(a_l)}\,, \tag{80}$$

since by definition, $p(c) = a_l$ for $c \in \sigma(a_l)$. We get

$$\begin{aligned}
\alpha_0 \prod_{i=1}^{k} \underline{i} &= \frac{1}{\mu_a} \prod_{i=1}^{k} \frac{\lambda_{a_{i-1}}}{\mu_{a_i}} \left(1 + \sum_{c \in \sigma(a)} lc\, t_c \right) \\
&= \frac{1}{\mu_{a_k} \pi(a_k)} \left(\pi(a) + \sum_{c \in \sigma(a)} \pi(B_c) \right) = \frac{\pi(B_a)}{\mu_{a_k} \pi(a_k)}\,,
\end{aligned} \tag{81}$$

and

$$\begin{aligned}
\alpha_l \prod_{i=l+1}^{k} \underline{i} &= \frac{1}{\mu_{a_l}} \prod_{i=l+1}^{k} \frac{\lambda_{a_{i-1}}}{\mu_{a_i}} \left(1 + \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} lc\, t_c \right) \\
&= \frac{1}{\mu_{a_k} \pi(a_k)} \left(\pi(a_l) + \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \pi(B_c) \right).
\end{aligned} \tag{82}$$

Thus

$$D_{a_j} = \sum_{k=j}^{n-1} \frac{1}{\mu_{a_k} \pi(a_k)} \left(\sum_{l=1}^{k} \left(\pi(a_l) + \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \pi(B_c) \right) + \pi(B_a) \right), \tag{83}$$

and equation (31) is proven, using the decomposition (6) of B_{a_k}.

Proof of (32): Suppose that $d(a) \geq j > n$.

The idea is now to iterate the difference equation in the other direction.

We write D_{a_j} as

$$D_{a_j} = D_{a_{d(a)}} - \sum_{k=j+1}^{d(a)} \left(D_{a_k} - D_{a_{k-1}} \right) , \tag{84}$$

which with the boundary condition (68) gives

$$D_{a_j} = \sum_{k=n+1}^{j} \left(D_{a_k} - D_{a_{k-1}} \right) . \tag{85}$$

We rewrite (69) as

$$\lambda_{a_{k-1}} \left(D_{a_k} - D_{a_{k-1}} \right) = \mu_{a_k} \left(D_{a_{k+1}} - D_{a_k} \right) + \left(1 + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc\,t_c \right), \tag{86}$$

which gives

$$D_{a_k} - D_{a_{k-1}} = \widetilde{\beta}_k \left(D_{a_{k+1}} - D_{a_k} \right) + \widetilde{\alpha}_k , \tag{87}$$

with $\widetilde{\alpha}_k := \dfrac{1}{\lambda_{a_{k-1}}} \left(1 + \sum\limits_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc\,t_c \right)$, and $\widetilde{\beta}_k := \dfrac{\mu_{a_k}}{\lambda_{a_{k-1}}}$ for $n < k < d(a)$.

Iteration up to $k = d(a)$ (recall $a_{d(a)} = 0$), yields

$$D_{a_k} - D_{a_{k-1}} = \prod_{i=k}^{d(a)-1} \widetilde{\beta}_i \left(D_{a_{d(a)}} - D_{a_{d(a)-1}} \right) + \sum_{l=k}^{d(a)-1} \widetilde{\alpha}_l \prod_{i=k}^{l-1} \widetilde{\beta}_i . \tag{88}$$

And by the boundary condition (71)

$$D_{a_{d(a)}} - D_{a_{d(a)-1}} = \frac{1}{\lambda_{a_{d(a)-1}}} \left(1 + \sum_{\substack{c \in \sigma(a_{d(a)}) \\ c \neq a_{d(a)-1}}} lc\,t_c \right) =: \widetilde{\alpha}_{d(a)} . \tag{89}$$

Thus

$$D_{a_k} - D_{a_{k-1}} = \sum_{l=k}^{d(a)} \widetilde{\alpha}_l \prod_{i=k}^{l-1} \widetilde{\beta}_i , \tag{90}$$

and then

$$D_{a_j} = \sum_{k=n+1}^{j} \sum_{l=k}^{d(a)} \widetilde{\alpha}_l \prod_{i=k}^{l-1} \widetilde{\beta}_i . \tag{91}$$

As above this can be expressed in terms of the i.p.m. π, we get

$$\begin{aligned} D_{a_j} &= \sum_{k=n+1}^{j} \frac{1}{\lambda_{a_{k-1}} \pi(a_k)} \sum_{l=k}^{d(a)} \left(\pi(a_l) + \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \pi(B_c) \right) \\ &= \sum_{k=n+1}^{j} \frac{1 - \pi(B_{a_{k-1}})}{\lambda_{a_{k-1}} \pi(a_k)} \\ &= \sum_{k=n}^{j-1} \frac{1 - \pi(B_{a_k})}{\mu_{a_k} \pi(a_k)} . \end{aligned} \tag{92}$$

The second equality is due to the decomposition (7), and the last one is obtained using reversibility. Finally, equation (32) is proven by inverting j and n in this last formula.

$\square$

Proof of Proposition 3.4

The proof follows the same lines as those of proposition 3.3.

We fix $a \in J$ and $0 \leq n \leq d(a)$. The application of (40) for $F(T) = T^2$ together with (9) gives

$$\mathbb{E}[T^2_{a_k \to a_n}] = \mu_{a_k} \mathbb{E}[T^2_{a_{k+1} \to a_n}] + \lambda_{a_{k-1}} \mathbb{E}[T^2_{a_{k-1} \to a_n}] + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc \, \mathbb{E}[T^2_{c \to a_n}] + 1 \tag{93}$$

$$+ 2 \left(\mu_{a_k} \mathbb{E}[T_{a_{k+1} \to a_n}] + \lambda_{a_{k-1}} \mathbb{E}[T_{a_{k-1} \to a_n}] + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc \, \mathbb{E}[T_{c \to a_n}] \right),$$

and by (65), the second line is equal to $2 \, \mathbb{E}[T_{a_k \to a_n}] - 2$.

We denote by Δ^2_x the second moment of $T_{x \to a_n}$ and keep the notation D_x for its expectation:

$$\Delta^2_x = \mathbb{E}[T^2_{x \to a_n}], \quad \text{and} \quad D_x = \mathbb{E}[T_{x \to a_n}] \tag{94}$$

Thus for each $0 < k < d(a)$, $k \neq n$, we have

$$\Delta^2_{a_k} = \mu_{a_k} \Delta^2_{a_{k+1}} + \lambda_{a_{k-1}} \Delta^2_{a_{k-1}} + \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc \, \Delta^2_c + 2 \, D_{a_k} - 1, \tag{95}$$

and the same computations in sites $a_0 = a$ and $a_{d(a)} = 0$, gives the boundary conditions

$$\Delta^2_{a_0} := \mathbb{E}[T^2_{a \to a_n}] = \mu_a \Delta^2_{a_1} + \sum_{c \in \sigma(a)} lc \, \Delta^2_c + 2 \, D_a - 1, \tag{96}$$

$$\Delta^2_{a_{d(a)}} := \mathbb{E}[T^2_{0 \to a_n}] = \lambda_{a_{d(a)-1}} \Delta^2_{a_{d(a)-1}} + \sum_{\substack{c \in \sigma(0) \\ c \neq a_{d(a)-1}}} lc \, \Delta^2_c + 2 \, D_0 - 1, \tag{97}$$

and for $k = n$

$$\Delta^2_{a_n} := \mathbb{E}[T^2_{a_n \to a_n}] = 0. \tag{98}$$

As in the proof of Proposition 3.3, we can decompose $T_{c \to a_n} = T_{c \to p(c)} + T_{a_k \to a_n}$, for $c \in \sigma(a_k)$, $c \neq a_{k-1}$, and use (9) to get the analogous difference equation of (69) for second moments

$$(\mu_{a_k} + \lambda_{a_{k-1}}) \Delta^2_{a_k} = \mu_{a_k} \Delta^2_{a_{k+1}} + \lambda_{a_{k-1}} \Delta^2_{a_{k-1}} \tag{99}$$

$$+ \sum_{\substack{c \in \sigma(a_k) \\ c \neq a_{k-1}}} lc \left(\mathbb{E}[T^2_{c \to p(c)}] + 2 \, t_c \, D_{a_k} \right) + 2 \, D_{a_k} - 1.$$

Proof of (34): When $0 \leq j < n$, following the proof of (31) we get

$$\Delta^2_{a_j} = \sum_{k=j}^{n-1} \sum_{l=0}^{k} \alpha_l \prod_{i=l+1}^{k} \underline{\mathrm{i}}, \tag{100}$$

with $\underline{\mathrm{i}} := \dfrac{\lambda_{a_{i-1}}}{\mu_{a_i}}$, $\alpha_l := \dfrac{1}{\mu_{a_l}} \left(\sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} lc \left(\mathbb{E}[T^2_{c \to p(c)}] + 2 \, t_c \, D_{a_l} \right) + 2 \, D_{a_l} - 1 \right)$ for $0 < i, l < n$,

and $\alpha_0 := \dfrac{1}{\mu_a} \left(\sum_{c \in \sigma(a)} lc \left(\mathbb{E}[T^2_{c \to p(c)}] + 2 \, t_c \, D_a \right) + 2 \, D_a - 1 \right)$.

In order to get expression (34), we have now to express these quantities in terms of i.p.m. π. Let us start with the following lemma:

Lemma 4.4.

$$\alpha_l \prod_{i=l+1}^{k} \dot{i} \;=\; \frac{1}{\mu_{a_k}\,\pi\,(a_k)}\left(\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a_l}-1\Big)\Big(\pi\,(B_{a_l})-\pi\,(B_{a_{l-1}})\Big)\right) \tag{101}$$

$$\alpha_0 \prod_{i=1}^{k} \dot{i} \;=\; \frac{1}{\mu_{a_k}\,\pi\,(a_k)}\left(\sum_{c\in\sigma(a)}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a}-1\Big)\pi\,(B_a)\right) \tag{102}$$

Proof of Lemma 4.4

Let $0 < l < n$. We recall that by (79), we have

$$\frac{1}{\mu_{a_l}}\prod_{i=l+1}^{k}\frac{\lambda_{a_{i-1}}}{\mu_{a_i}} \;=\; \frac{\pi\,(a_l)}{\mu_{a_k}\,\pi\,(a_k)}. \tag{103}$$

It remains to compute the term in the parenthesis in the definition of α_l.

Using reversibility, $\dfrac{lc}{\mu_c\,\pi\,(c)} = \dfrac{1}{\pi\,(a_l)}$ for $c \in \sigma(a_l)$, and then by equation (28). we have

$$\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}} lc\,\Big(\mathbb{E}\big[T^2_{c\to p(c)}\big] + 2\,t_c\,D_{a_l}\Big) + 2\,D_{a_l} - 1$$

$$=\quad \frac{1}{\pi\,(a_l)}\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a_l}-1\Big)\Big(1+\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}} lc\,t_c\Big) \tag{104}$$

$$=\quad \frac{1}{\pi\,(a_l)}\left(\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a_l}-1\Big)\Big(\pi\,(B_{a_l})-\pi\,(B_{a_{l-1}})\Big)\right),$$

since by (13)

$$\Big(1+\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}} lc\,t_c\Big) \;=\; \frac{1}{\pi\,(a_l)}\left(\pi\,(a_l)+\sum_{\substack{c\in\sigma(a_l)\\ c\neq a_{l-1}}}\pi\,(B_c)\right) \tag{105}$$

$$=\quad \frac{1}{\pi\,(a_l)}\Big(\pi\,(B_{a_l})-\pi\,(B_{a_{l-1}})\Big).$$

Equation (101) is the consequence of (103) and (104).

For the proof of (102) we use similar arguments for $l = 0$:

$$\frac{1}{\mu_a}\prod_{i=1}^{k}\frac{\lambda_{a_{i-1}}}{\mu_{a_i}} \;=\; \frac{\pi\,(a)}{\mu_{a_k}\,\pi\,(a_k)}, \tag{106}$$

and

$$\sum_{c\in\sigma(a)} lc\,\Big(\mathbb{E}\big[T^2_{c\to p(c)}\big] + 2\,t_c\,D_{a}\Big) + 2\,D_{a} - 1$$

$$=\quad \frac{1}{\pi\,(a)}\sum_{c\in\sigma(a)}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a}-1\Big)\Big(1+\sum_{c\in\sigma(a)} lc\,t_c\Big) \tag{107}$$

$$=\quad \frac{1}{\pi\,(a)}\left(\sum_{c\in\sigma(a)}\sum_{b\in B_c}\frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \Big(2\,D_{a}-1\Big)\pi\,(B_a)\right).$$

$$\square$$

We now have

$$\Delta_{a_j}^2 = \sum_{k=j}^{n-1} \frac{1}{\mu_{a_k} \pi(a_k)} \left(\sum_{l=1}^{k} \Gamma_l + \Gamma_0 \right), \tag{108}$$

with

$$\Gamma_l := \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \sum_{b \in B_c} \frac{2\pi(B_b)^2}{\mu_b \pi(b)} + \left(2 D_{a_l} - 1 \right) \left(\pi(B_{a_l}) - \pi(B_{a_{l-1}}) \right) \tag{109}$$

and

$$\Gamma_0 := \sum_{c \in \sigma(a)} \sum_{b \in B_c} \frac{2\pi(B_b)^2}{\mu_b \pi(b)} + \left(2 D_a - 1 \right) \pi(B_a). \tag{110}$$

Summing over l, the second term in (109), we get

$$\sum_{l=1}^{k} \left(2 D_{a_l} - 1 \right) \left(\pi(B_{a_l}) - \pi(B_{a_{l-1}}) \right)$$

$$= \sum_{l=1}^{k} \left(2 D_{a_l} - 1 \right) \pi(B_{a_l}) - \sum_{l=1}^{k} \left(2 D_{a_l} - 1 \right) \pi(B_{a_{l-1}})$$

$$= \sum_{l=1}^{k} \left(2 \left(t_{a_l} + D_{a_{l+1}} \right) - 1 \right) \pi(B_{a_l}) - \sum_{l=0}^{k-1} \left(2 D_{a_{l+1}} - 1 \right) \pi(B_{a_l}) \tag{111}$$

$$= \sum_{l=1}^{k} 2 t_{a_l} \pi(B_{a_l}) + \sum_{l=1}^{k} \left(2 D_{a_{l+1}} - 1 \right) \pi(B_{a_l}) - \sum_{l=0}^{k-1} \left(2 D_{a_{l+1}} - 1 \right) \pi(B_{a_l})$$

$$= \sum_{l=1}^{k} \frac{2\pi(B_{a_l})^2}{\mu_{a_l} \pi(a_l)} + \left(2 D_{a_{k+1}} - 1 \right) \pi(B_{a_k}) - \left(2 D_{a_1} - 1 \right) \pi(B_a),$$

where we have used that for $l < n$, we have $T_{a_l \to a_n} = T_{a_l \to a_{l+1}} + T_{a_{l+1} \to a_n}$.
And the boundary condition (70) gives

$$\pi(B_a) \left(\left(2 D_a - 1 \right) - \left(2 D_{a_1} - 1 \right) \right)$$

$$= 2\pi(B_a) \left(D_a - D_{a_1} \right)$$

$$= \frac{2\pi(B_a)}{\mu_a} \left(1 + \sum_{c \in \sigma(a)} lc t_c \right)$$

$$= \frac{2\pi(B_a)}{\mu_a \pi(a)} \left(\pi(a) + \sum_{c \in \sigma(a)} \pi(B_c) \right)$$

$$= \frac{2\pi(B_a)^2}{\mu_a \pi(a)}.$$

Thus

$$\Gamma_0 - \left(2 D_{a_1} - 1 \right) \pi(B_a) = \sum_{c \in \sigma(a)} \sum_{b \in B_c} \frac{2\pi(B_b)^2}{\mu_b \pi(b)} + \frac{2\pi(B_a)^2}{\mu_a \pi(a)} = \sum_{b \in B_a} \frac{2\pi(B_b)^2}{\mu_b \pi(b)}, \tag{112}$$

since by definition $B_a = \{a\} \bigcup_{c \in \sigma(a)} B_c$.

Hence

$$\sum_{l=1}^{k} \Gamma_l + \Gamma_0 \;=\; \sum_{l=1}^{k} \left(\sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \sum_{b \in B_c} \frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \frac{2\,\pi\,(B_{a_l})^2}{\mu_{a_l}\,\pi\,(a_l)} \right) + \sum_{b \in B_a} \frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} \tag{113}$$

$$+ \left(2\,D_{a_{k+1}} - 1 \right) \pi\,(B_{a_k}) \tag{114}$$

$$=\; \sum_{b \in B_{a_k}} \frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + \left(2\,D_{a_{k+1}} - 1 \right) \pi\,(B_{a_k}) \,, \tag{115}$$

where the second equality is due to the decomposition (6) of B_{a_k}.
Finally,

$$\Delta_{a_j}^2 \;=\; \sum_{k=j}^{n-1} \frac{1}{\mu_{a_k}\,\pi\,(a_k)} \left(\sum_{b \in B_{a_k}} \frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} + 2\,\pi\,(B_{a_k})\,D_{a_{k+1}} - \pi\,(B_{a_k}) \right), \tag{116}$$

and equation (34) is obtained using (31).

Proof of (35): Suppose that $d(a) \geq j > n$.
As in the proof of (32), from the difference equation (99), we obtain

$$\Delta_{a_j}^2 \;=\; \sum_{k=n+1}^{j} \sum_{l=k}^{d(a)} \widetilde{\alpha}_l \prod_{i=k}^{l-1} \widetilde{\beta}_i \,, \tag{117}$$

with $\widetilde{\alpha}_l := \dfrac{1}{\lambda_{a_{l-1}}} \left(\displaystyle\sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} l c \left(\mathbb{E}\big[T_{c \to p(c)}^2\big] + 2\,t_c\,D_{a_l} \right) + 2\,D_{a_l} - 1 \right)$ for $n < l \leq d(a)$,

and $\widetilde{\beta}_i := \dfrac{\mu_{a_i}}{\lambda_{a_{i-1}}}$ for $n < i < d(a)$.

By the same arguments as those of the proof of Lemma 4.4, we get $\widetilde{\alpha}_l \displaystyle\prod_{i=k}^{l-1} \widetilde{\beta}_i \;=\; \dfrac{1}{\lambda_{a_{k-1}}\,\pi\,(a_k)}\,\Gamma_l$, where Γ_l is defined by equation (109). And thus, using reversibility

$$\Delta_{a_j}^2 \;=\; \sum_{k=n+1}^{j} \frac{1}{\lambda_{a_{k-1}}\,\pi\,(a_k)} \sum_{l=k}^{d(a)} \Gamma_l \;=\; \sum_{k=n}^{j-1} \frac{1}{\mu_{a_k}\,\pi\,(a_k)} \sum_{l=k+1}^{d(a)} \Gamma_l \,. \tag{118}$$

It remains to compute $\displaystyle\sum_{l=k+1}^{d(a)} \Gamma_l$. First, by definition (8), we have

$$\sum_{l=k+1}^{d(a)} \sum_{\substack{c \in \sigma(a_l) \\ c \neq a_{l-1}}} \sum_{b \in B_c} \frac{2\,\pi\,(B_b)^2}{\mu_b\,\pi\,(b)} \;=\; \sum_{c \in C_{a_k}} \frac{2\,\pi\,(B_c)^2}{\mu_c\,\pi\,(c)} \,. \tag{119}$$

For the sum of the second term in (109), we use a computation similar to (111) with the following decomposition: for $l > n$, we have $T_{a_{l+1} \to a_n} = T_{a_{l+1} \to a_l} + T_{a_l \to a_n}$, and thus

$$\sum_{l=k+1}^{d(a)} \left(2\,D_{a_l} - 1 \right) \left(\pi\,(B_{a_l}) - \pi\,(B_{a_{l-1}}) \right)$$

$$=\; \sum_{l=k+1}^{d(a)} \left(2\,D_{a_l} - 1 \right) \pi\,(B_{a_l}) - \sum_{l=k}^{d(a)-1} \left(2 \left(\mathbb{E}\big[T_{a_{l+1} \to a_l}\big] + D_{a_l} \right) - 1 \right) \pi\,(B_{a_l}) \tag{120}$$

$$=\; \left(2\,D_{a_{d(a)}} - 1 \right) - \left(2\,D_{a_k} - 1 \right) \pi\,(B_{a_k}) - 2 \sum_{l=k}^{d(a)-1} \mathbb{E}\big[T_{a_{l+1} \to a_l}\big]\,\pi\,(B_{a_l}) \,,$$

21

since by definition $B_{a_{d(a)}} = B_0 = I$.

We now write $D_{a_{d(a)}} := \mathbb{E}\big[T_{a_{d(a)} \to a_n}\big]$ as $\mathbb{E}\big[T_{a_{d(a)} \to a_k}\big] + D_{a_k}$, and since by (32), we have

$$\mathbb{E}\big[T_{a_{d(a)} \to a_k}\big] = \sum_{l=k}^{d(a)-1} \mathbb{E}\big[T_{a_{l+1} \to a_l}\big] = \sum_{l=k}^{d(a)-1} \frac{1 - \pi(B_{a_l})}{\mu_{a_l} \, \pi(a_l)}, \qquad (121)$$

we get

$$\sum_{l=k+1}^{d(a)} \Big(2\, D_{a_l} - 1 \Big) \Big(\pi(B_{a_l}) - \pi(B_{a_{l-1}}) \Big)$$

$$= \Big(2\, D_{a_k} - 1 \Big) \Big(1 - \pi(B_{a_k}) \Big) + 2 \sum_{l=k}^{d(a)-1} \frac{\Big(1 - \pi(B_{a_l}) \Big)^2}{\mu_{a_l} \, \pi(a_l)} \qquad (122)$$

$$= \Big(2\, D_{a_k} - 1 \Big) \pi\big(\overline{B}_{a_k}\big) + 2 \sum_{l=k}^{d(a)-1} \frac{\pi\big(\overline{B}_{a_l}\big)^2}{\mu_{a_l} \, \pi(a_l)}.$$

Finally

$$\Delta_{a_j}^2 = \sum_{k=n}^{j-1} \frac{1}{\mu_{a_k} \, \pi(a_k)} \left(\sum_{c \in C_{a_k}} \frac{2\, \pi(B_c)^2}{\mu_c \, \pi(c)} + 2 \sum_{l=k}^{d(a)-1} \frac{\pi\big(\overline{B}_{a_l}\big)^2}{\mu_{a_l} \, \pi(a_l)} + \Big(2\, D_{a_k} - 1 \Big) \pi\big(\overline{B}_{a_k}\big) \right), \quad (123)$$

and equation (35) is obtained by inverting j and n, using expression (32) for $\mathbb{E}\big[T_{a_n \to a_j}\big]$, and since by definition (4), the path $\ell(a_k)$ is equal to $\cup_{l=k}^{d(a)-1}\{a_l\}$.

$\square$